DOCUMENT RESUME

ED 058 156 SP 005 455

AUTHOR Throne, John M.
TITLE Does Operant Conditioning Make Functional Distinctions Between Teaching and Evaluation Obsolete?
PUB DATE 71
NOTE 10p.; Based on a paper presented to the meeting of the American Psychological Assn., Washington, D.C., September 19, 1971

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Educational Responsibility; *Effective Teaching; *Learning Processes; *Operant Conditioning

ABSTRACT

The differences between the authors of the Coleman report and their critics make clear why the effectiveness of schools will always remain controversial as long as inferential statistics are employed to determine it. Controversy is inevitable when measurement requires the satisfaction of assumptions antithetical to the teaching processes through which the effects of schools must be produced. Operant conditioning, by providing means for determining independent-variable effectiveness and dependent-variable effects in a single set of operants, lets such counterproductive assumptions be set aside. (Author)

EDU 58156

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE O EDUCATION POSITION OR POLICY

REVISED

# DOES OPERANT CONDITIONING MAKE FUNCTIONAL DISTINCTIONS BETWEEN TEACHING AND EVALUATION OBSOLETE?[1]

John M. Throne, Ph.D.

University of Kansas

SP 005455

[1]Based on a paper presented to the meeting of the American Psychological Association, Washington, September, 1971.

ABSTRACT

The differences between the authors of the Coleman report and their critics make clear why the effectiveness of schools will always remain controversial as long as inferential statistics are employed to determine it. Controversy is inevitable when measurement requires the satisfaction of assumptions antithetical to the teaching processes through which the effects of schools must be produced. Operant conditioning, by providing means for determining independent-variable effectiveness and dependent-variable effects in a single set of operations, lets such counterproductive assumptions be set aside.

In 1966 James S. Coleman and several collaborators (Coleman, et al, 1966 -- hereinafter referred to as Coleman) published a report charging, in effect, that the American public school is irrelevant. According to Coleman, "schools bring little to bear on a child's academic achievement that is independent of his background and general social context" (p. 325). Coleman asserted that "there is a relatively small amount of school-to-school variation that is not accounted for by differences in family background, indicating the small independent effect of variations in school facilities, curriculum, and staff upon achievement" (p. 325).

## Are Schools Irrelevant?

Needless to say, Coleman's educational colleagues were shocked. For example, Egon G. Guba and David L. Clark (Guba and Clark, 1967) found this conclusion "incredible on its face" (p. 95). They declared it contradicts "tens of thousands of independent observations which have noted differences in student behavior in different educational settings, even taking into account differences in family background" (p. 95). But one independent observation ($x$) does not validate another ($y$), except that correlation of the two, with $p$ probability under $q$ conditions, is arbitrarily defined as $y$'s criterion. Guba and Clark appear to have confused validity with reliability. In any event, Coleman did not deny that individual pupils behave differently in different settings, only that a functional relationship between settings and achievement has been proved.

A relationship is functional insofar as one variable, $a$, controls another variable, $b$; in which case the latter is said to be a function of the former. Guba and Clark might have pointed to Coleman's questionable

logic in concluding, in the absence of proof that achievement is a function of scholastic variables, that extrascholastic variables necessarily cause whatever changes in achievement occur. At issue is the logicality of conclusions based on failed hypotheses. Failure to confirm an hypothesis in no way accords presumptive proof to its antithesis, nor does it preclude the possibility of opposite results if alternative conditions are introduced. However, they put forth a logically questionable conclusion of their own. They contended that "the typical finding of 'no significant difference' strongly implies that no new technique, no new practice, no new methodology invented by experienced educators is ever any better or any worse than what has been going on all along" (p. 9). Guba and Clark found this conclusion "statistically incredible" (p. 9). They exaggerated, no doubt for rhetorical purposes, both the success and unsuccess revealed by studies of effectiveness of schools. But hyperbole aside, surely the frequency of negative findings is insufficient grounds for reasoning that those findings must in fact be positive. Failure to reject the null hypothesis casts doubt on the experimental hypothesis, not the null; failure to reject the null again and again does not reduce that doubt by one iota.

## Assumptions of Inferential Statistics: Unfeasible or Worse?

But the positions held by Coleman on the one hand and Guba and Clark on the other are vulnerable at a more fundamental level. The conclusions of both Coleman and Guba and Clark were based on studies employing the evaluative processes of inferential statistics, which require that the presence or absence of functional relationships between scholastic or extra-scholastic and achievement variables be deduced. But deducing that func-

relationships are present or absent falls short of demonstrating they do or do not exist in fact. Deduction is at best a prelude to demonstration, at worst a tautological exercise beyond proof and disproof alike.

Guba and Clark criticized Coleman on the secondary issue of which inferences may be validly arrived at on the basis of manipulation of semantic signs: the numerical symbols constituting the manipulanda of inferential statistics. But the primary issue is, How are the phenomenal-significate responses of students on which semantic-sign manipulation is based -- responses constituting the manipulanda of teaching--effectively produced by schools? It is only through production of such responses by schools that functional relationships between schools and scholastic achievement may be demonstrably made known. The same applies to extrascholastic and achievement variables; the former must produce the latter before functional relationships between them may be known to operate.

In Guba and Clark's judgment, the negative findings which Coleman's unfavorable judgment of schools is based on are cause by misapplication of inferential statistics despite the fact that the assumptions requisite to their employment, in the typical case, remain unmet. "All of these assumptions," Guba and Clark observed, "are in some particulars unrealistic for education" (p. 107). But only belatedly did they begin to approach the heart of the matter: the inappropriateness of inferential statistics in evaluating scholastic effectiveness precisely to the extent that these assumptions have been met. Almost in passing, they remarked that treatment invariance is not only quite difficult to achieve but may be undesirable, since the treatment may be one that could profit from continuous improvement

even while being tested (p. 107). The desirability of treatment variance during testing most certainly applies (as Guba and Clark made clear) to education. But this in turn requires (as they did not make clear) an evaluation system that is indistinguishable from teaching; a framework within whose bounds may be found a single set of operations for determining the impact of educational practices in both of the senses, teaching and evaluation, that the term determination implies (Throne, 1970). Through operations called teaching (or training, etc.), an educator may determine, in the sense of produce, the level of achievement a child attains. The educator may also determine, in the sense of measure, that selfsame level of achievement. However, if measurement is undertaken through an active process of response manipulation to criterion, rather than (as in the case of inferential statistics) a passive one of comparison of obtained results against a theoretical expectancy (i.e., null hypothesis testing), all operational distinctions between teaching and evaluation may be dissolved (Throne, 1971a, b).

In effect, if teaching implies the production rather than measurement of behavior, then, given the manipulative processes intrinsic to teaching, the assumptions necessary for evaluating teaching through inferential statistics must not be met. The demands of teaching, not those of evaluation antithetical to teaching, must dictate the arrangements under which evaluation data on teaching are obtained (Throne, 1971a, b). To measure the effectiveness of teaching under the restrictions on teaching which assumptions aside from teaching impose, is not to measure teaching effectiveness at all. To label inferential statistics as inappropriate

in evaluating teaching effectiveness on grounds of unfeasibility is to miss the point. They are inappropriate because the assumptions making their operations valid are ipso facto antithetical to the valid operations of teaching.

## Operant Conditioning

That different operations are implicit in the terms teaching and evaluation is an article of faith in education. It is therefore not surprising that educators seem unaware that a basis for producing teaching effects and measuring teaching effectiveness through a single set of operations has long been available in the strategies and tactics of the operant conditioning model developed by B.F. Skinner (e.g., Skinner 1938, 1965, 1968). (Two brand new books on evaluation containing contributions by leading educational specialists include not a single reference to Skinner. They are: Research in Teacher Education: A Symposium, B. Othanel Smith, editor, Englewood Cliffs, N.J.: Prentice-Hall, 1971, 166 pp.; and Educational Evaluation and Decision Making, Phi Delta Kappa National Study Commission on Education, Itasca, Ill.: F.E. Peacock, 1971, 368 pp.) Fundamentally, these operations entail manipulation of the consequential stimuli that responses encounter, sufficient to produce dependent-variable responses to criterion. What has been referred to as the principle of consequential determinism (Throne, 1970) is the keystone on which these strategies and tactics have been built.

According to the principle of consequential determinism: "Behavior is a function of its consequences" (Skinner, 1938, p. 63). That is, the probability of a member of a behavioral class occurring is a function of the consequences that

other members of the class encounter. If consequences are reinforcing, the probability of occurrence will increase; if extinguishing or aversive, it will decrease. If consequences are neutral, it will remain unchanged. It follows that the presentation or withdrawal (or withholding) of whichever consequences demonstrably increase or decrease the probability of criterion behavior, or successive approximations of it, may be made <u>contingent</u> on such increases or decreases, as the case may be. It further follows that if the effects of teaching are determined (in the sense of produced) on a contingent basis, then the effectiveness of consequences may be simultaneously determined (in the sense of measured) by those self-same effects. Neither the success of teaching nor its failure need be determined (in either sense) independently of teaching. For example, if a subject is taught to write his name, count to five, or spell 10 words, the effectiveness of the operations employed is revealed in the effects obtained. Ineffectiveness is shown by un-effects. In either case, evaluative operations distinct from teaching are unnecessary. <u>Teaching and evaluation thus coalesce</u> (Throne, 1971a, b).

The principle of consequential determinism is the key, therefore, to the problem raised by the inappropriateness under any circumstances of inferential statistics in evaluating teaching. Insofar as an educator succeeds in introducing or withdrawing (or withholding) consequential stimuli sufficient to determine (produce) criterion responses, he determines (measures) by the outcome the effectiveness of the stimuli employed.

Precisely <u>which</u> components of effective stimuli evoke obtained response effects can be determined (produced and measured) through a differential reinforcement strategy in which all stimulus components other than those

in question are controlled (so-called "component analysis"). Thus the issue of stimulus necessity is transformed into one of stimulus sufficiency. Of course, the magnitude, longevity, and generalizability of effects (and certain "side-effects") of consequential stimuli will vary, depending on the reinforcement history of the subjects on whom the stimuli impinge. These effects are not immutable; they may be altered by presenting or withdrawing (or withholding) other consequences with different reinforcement values. For example, if contingent teacher attention fails to reinforce academic achievement, maybe special privileges will (e.g., opportunity to leave classroom early). Also, effects of consequences can be reversed. For example, teacher attention may be made reinforcing by pairing it with a stimulus that already is (e.g., classroom recitation). As the reinforcement value of teacher attention subsequently increases, it may be used to reinforce criterion behavior (academic achievement), not possible before.

## Conclusion

The differences between critics of teaching like Coleman and of evaluation like Guba and Clark make clear why school effectiveness will always remain controversial as long as inferential statistics are employed to measure it. Controversy is inevitable when evaluation demands, feasibly or not, that assumptions antithetical to teaching be fulfilled. This issue is obviated by operant conditioning. Operant conditioning measures teaching effectiveness by producing achievement effects, guaranteeing they reflect the operations of teaching only, instead of evaluative operations confounded with and hobbling teaching. In a word, operant conditioning makes functional distinctions between teaching and evaluation obsolete.

References

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weingeld, F.D., and York, R.L. Equality of educational opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.

Guba, E.G. and Clark, D.L. A proposal for the National Institute for the Study of Educational Change. NISEC, School of Education, Indiana University, Bloomington, Indiana, 1967.

Skinner, B.F. Behavior of organisms. New York: Appleton-Century-Crofts, 1938.

Skinner, B.F. Science and human behavior. New York: Free Press, 1965.

Skinner, B.F. The technology of teaching. New York: Appleton-Century-Crofts, 1968.

Throne, J.M. A radical behaviorist approach to diagnosis in mental retardation. Mental Retardation, 1970, 8(3), 2-5.

Throne, J.M. Inappropriateness of inferential and insufficiency of descriptive statistics in evaluating effectiveness of teaching: The problem and a solution. Presented to the meeting of the American Education Research Association, New York City, February, 1971. (a)

Throne, J.M. Statistical analysis in the evaluation of teaching: Is it necessary? Educational Researcher, March, 1971, p. 4-6. (b)